

=====

Sequence Listing was accepted.

See attached Validation Report.

If you need help call the Patent Electronic Business Center at (866)
217-9197 (toll free).

Reviewer: markspencer

Timestamp: [year=2008; month=4; day=23; hr=14; min=8; sec=56; ms=527;]

=====

Application No: 10620099 Version No: 3.0

Input Set:

Output Set:

Started: 2008-04-09 15:16:12.876
Finished: 2008-04-09 15:16:13.531
Elapsed: 0 hr(s) 0 min(s) 0 sec(s) 655 ms
Total Warnings: 21
Total Errors: 1
No. of SeqIDs Defined: 23
Actual SeqID Count: 23

Error code	Error Description
W 213	Artificial or Unknown found in <213> in SEQ ID (3)
W 213	Artificial or Unknown found in <213> in SEQ ID (4)
W 213	Artificial or Unknown found in <213> in SEQ ID (5)
E 257	Invalid sequence data feature in <221> in SEQ ID (5)
W 251	Found intentionally skipped sequence in SEQID (6)
W 251	Found intentionally skipped sequence in SEQID (7)
W 251	Found intentionally skipped sequence in SEQID (8)
W 251	Found intentionally skipped sequence in SEQID (9)
W 251	Found intentionally skipped sequence in SEQID (10)
W 251	Found intentionally skipped sequence in SEQID (11)
W 251	Found intentionally skipped sequence in SEQID (12)
W 251	Found intentionally skipped sequence in SEQID (13)
W 251	Found intentionally skipped sequence in SEQID (14)
W 251	Found intentionally skipped sequence in SEQID (15)
W 251	Found intentionally skipped sequence in SEQID (16)
W 251	Found intentionally skipped sequence in SEQID (17)
W 251	Found intentionally skipped sequence in SEQID (18)
W 251	Found intentionally skipped sequence in SEQID (19)
W 213	Artificial or Unknown found in <213> in SEQ ID (20)
W 213	Artificial or Unknown found in <213> in SEQ ID (21)

Input Set:

Output Set:

Started: 2008-04-09 15:16:12.876
Finished: 2008-04-09 15:16:13.531
Elapsed: 0 hr(s) 0 min(s) 0 sec(s) 655 ms
Total Warnings: 21
Total Errors: 1
No. of SeqIDs Defined: 23
Actual SeqID Count: 23

Error code	Error Description
W 213	Artificial or Unknown found in <213> in SEQ ID (22)
W 213	Artificial or Unknown found in <213> in SEQ ID (23)

SEQUENCE LISTING

<110> WACHTER, Rebekka M.

REMINGTON, S. James

<120> LONG WAVELENGTH ENGINEERED FLUORESCENT PROTEINS

<130> 026069-151480

<140> 10620099

<141> 2003-07-14

<150> US 09/575,847

<151> 2000-05-19

<150> US 08/974,737

<151> 1997-11-19

<150> US 08/911,825

<151> 1997-08-15

<150> US 08/706,408

<151> 1996-08-30

<150> US 60/024,050

<151> 1996-08-16

<160> 23

<170> PatentIn version 3.0

<210> 1

<211> 716

<212> DNA

<213> Aequorea victoria

<400> 1

```
atgagtaaag gagaagaact ttctactgca gttgtcccaa ttcttggtga attagatggt      60
gatgttaatg ggcacaaatt ttctgtcagt ggagaggggtg aaggtgatgt aacatacggg      120
aaacttaccc ttaaatttat ttgcactact ggaaaactac ctgttccatg gccaacactt      180
gtcactactt tctcttatgg tgttcaatgc ttttcaagat acccagatca tatgaaacgg      240
catgactttt tcaagagtgc catgcccgaa ggttatgtac agcaaagaac tatatttttc      300
aaagatgacg ggaactacaa gacacgtgct gaagtcaagt ttgaaggtga tacccttggt      360
aatagaatcg agttaaaggg tattgatttt aaagaagatg gaaacattct tggacataaa      420
ttggaataca actataactc acacaatgta tacatcatgg cagacaaaca aaagaatgga      480
atcaaagtta acttcaaaat tagacacaac attgaagatg gaagcgttca actagcagac      540
tattatcaac aaaatactcc aattctcgat ggccctgtcc ttttaccaga caaccattac      600
ctgtccacac aatctgccct ttcgaaagat cccaacgaaa agagagacca catggtcctt      660
cttgagtttg taacagctgc tgggattaca catggcatgg atgaactata caaata          716
```

<210> 2

<211> 238

<212> PRT

<213> Aequorea victoria

<400> 2

```
Met Ser Lys Gly Glu Glu Leu Phe Thr Ala Val Val Pro Ile Leu Val
1              5              10              15
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20              25              30
Gly Glu Gly Asp Val Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35              40              45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
50              55              60
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
65              70              75              80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Gln Arg
85              90              95
```

Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val
			100					105					110		
Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile
		115					120					125			
Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn
		130				135					140				
Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly
145					150					155					160
Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val
				165					170					175	
Gln	Leu	Ala	Asp	Tyr	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Leu	Asp	Gly	Pro
			180					185					190		
Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser
		195					200					205			
Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val
		210				215					220				
Thr	Ala	Ala	Gly	Ile	Thr	His	Gly	Met	Asp	Glu	Leu	Tyr	Lys		
225					230					235					

<210> 3

<211> 720

<212> DNA

<213> Artificial sequence

<220>

<223> DNA Sequence encoding engineered Aequorea-related fluorescent protein

<400> 3

atggtgagca	agggcgagga	gctgttcacc	ggggtggtgc	ccatcctggt	cgagctggac	60
ggcgacgtaa	acggccacaa	gttcagcgtg	tccggcgagg	gcgagggcga	tgccacctac	120
ggcaagctga	ccctgaagtt	catctgcacc	accggcaagc	tgcccgtgcc	ctggcccacc	180
ctcgtgacca	ccttcggcta	cggcgtgcag	tgcttcgccc	gctaccccga	ccacatgaag	240
cagcaggact	tcttcaagtc	cgccatgccc	gaaggetacg	tccaggagcg	caccatcttc	300
ttcaaggacg	acggcaacta	caagaccgcg	gccgaggtga	agttcgaggg	cgacaccctg	360
gtgaaccgca	tcgagctgaa	gggcatcgac	ttcaaggacg	acggcaacat	cctggggcac	420
aagctggagt	acaactacaa	cagccacaac	gtctatatca	tggccgacaa	gcagaagaac	480
ggcatcaagg	tgaacttcaa	gatccgccac	aacatcgagg	acggcagcgt	gcagcccgcc	540
gaccactacc	agcagaacac	ccccatcggc	gacggccccg	tgctgtgtgc	cgacaaccac	600
tacctgagct	accagtccgc	cctgagcaaa	gacccaacg	agaagcgcga	tcacatggtc	660
ctgctggagt	tcgtgaccgc	cgccgggatc	actcacggca	tggaagagct	gtacaagtaa	720

<210> 4

<211> 239

<212> PRT

<213> Artificial sequence

<220>

<223> Engineered Aequorea-related fluorescent protein

<400> 4

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu
1			5					10					15		
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly
		20					25					30			
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile
		35				40					45				
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr
	50				55					60					
Phe	Gly	Tyr	Gly	Val	Gln	Cys	Phe	Ala	Arg	Tyr	Pro	Asp	His	Met	Lys
65				70					75					80	

Gln Gln Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95
 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110
 Val Lys Phe Phe Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125
 Ile Asp Phe Lys Asp Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130 135 140
 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 145 150 155 160
 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
 165 170 175
 Val Gln Pro Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 180 185 190
 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
 195 200 205
 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210 215 220
 Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 225 230 235

<210> 5

<211> 228

<212> PRT

<213> Artificial sequence

<220>

<223> Engineered Aequorea-related fluorescent protein

<220>

<221> MOD_RES

<222> (64)..(66)

<223> The amide nitrogen of Gly 66 is cyclized onto the amide of Tyr
 65, the amide oxygen of Thr 64 has been removed, and the bond
 between the alpha and beta carbons of Tyr 65 is oxidized to form
 a conjugated GFP chromophore.

<400> 5

Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu
 1 5 10 15
 Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly
 20 25 30
 Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr
 35 40 45
 Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe Thr
 50 55 60
 Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg His
 65 70 75 80
 Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr
 85 90 95
 Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys
 100 105 110
 Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp
 115 120 125
 Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr
 130 135 140
 Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile

145		150		155		160									
Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val	Gln
				165					170					175	
Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	Pro	Val
			180					185					190		
Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser	Lys
		195					200				205				
Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val	Thr
	210					215				220					
Ala	Ala	Gly	Ile												
225															

<210> 6
 <400> 6
 000

<210> 7
 <400> 7
 000

<210> 8
 <400> 8
 000

<210> 9
 <400> 9
 000

<210> 10
 <400> 10
 000

<210> 11
 <400> 11
 000

<210> 12
 <400> 12
 000

<210> 13
 <400> 13
 000

<210> 14
 <400> 14
 000

<210> 15
 <400> 15
 000

<210> 16
 <400> 16
 000

<210> 17

<400> 17
000

<210> 18
<400> 18
000

<210> 19
<400> 19
000

<210> 20
<211> 5
<212> PRT
<213> Artificial sequence
<220>
<223> Synthetic localization sequence targeting the nucleus
<400> 20
Lys Lys Lys Arg Lys
1 5

<210> 21
<211> 26
<212> PRT
<213> Artificial sequence
<220>
<223> Synthetic localization sequence targeting the mitochondrion
<400> 21
Met Leu Arg Thr Ser Ser Leu Phe Thr Arg Arg Val Gln Pro Ser Leu
1 5 10 15
Phe Arg Asn Ile Leu Arg Leu Gln Ser Thr
20 25

<210> 22
<211> 4
<212> PRT
<213> Artificial sequence
<220>
<223> Synthetic localization sequence targeting the endoplasmic
reticulum
<400> 22
Lys Asp Glu Leu
1

<210> 23
<211> 37
<212> PRT
<213> Artificial sequence
<220>
<223> Synthetic His-tag amino acid sequence
<400> 23
Met Arg Gly Ser His His His His His Gly Met Ala Ser Met Thr
1 5 10 15
Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp
20 25 30
Pro Pro Ala Glu Phe
35

